Amendments to the Specification:

Please replace the paragraph on page 11 starting on line 6 with the following amended paragraph:

Microarrays can be used in a variety of ways. A preferred microarray contains nucleic acids and is used to analyze nucleic acid samples. Typically, a nucleic acid sample is prepared from appropriate source and labeled with a signal moiety, such as a fluorescent label. The sample is hybridized with the array under appropriate conditions. The arrays are washed or otherwise processed to remove non-hybridized sample nucleic acids. The hybridization is then evaluated by detecting the distribution of the label on the chip. The distribution of label may be detected by scanning the arrays to determine fluorescence intensity distribution. Typically, the hybridization of each probe is reflected by several pixel intensities. The raw intensity data may be stored in a gray scale pixel intensity file. The GATC™ Consortium has specified several file formats for storing array intensity data. The final software specification is available on the internet at ~~www.gatcconsortium.org~~ gatcconsortium.com and is incorporated herein by reference in its entirety. The pixel intensity files are usually large. For example, a GATC™ compatible image file may be approximately 50 Mb if there are about 5000 pixels on each of the horizontal and vertical axes and if a two byte integer is used for every pixel intensity. The pixels may be grouped into cells. (*See* GATC™ software specification). The probes in a cell are designed to have the same sequence; i.e., each cell is a probe area. A CEL file contains the statistics of a cell, e.g., the 75th percentile and standard deviation of intensities of pixels in a cell. The 50, 60, 70, 75 or 80th percentile of pixel intensity of a cell is often used as the intensity of the cell.

Please replace the paragraph on page 12 starting on line 4 with the following amended paragraph:

The Affymetrix® Analysis Data Model (AADM) is the relational database schema Affymetrix uses to store experiment results. It includes tables to support mapping, spotted arrays and expression results. Affymetrix publishes AADM to support open access to experiment information generated and managed by Affymetrix® software so that results may be filtered and mined with any compatible analysis tools. The AADM specification (Affymetrix, Santa Clara, CA, 2001) is incorporated herein by reference for all purposes. The specification is available at the Affymetrix website ~~http://www.affymetrix.com/support/aadm/aadm.html, last visited on 9/4/2001~~.

Please replace the paragraph on page 15 starting on line 13 with the following amended paragraph:

FIG. 3 shows an exemplary computer network that is suitable for executing the computer software of the invention. A computer workstation 302 is connected with and controls a probe array scanner 301. Probe intensities are acquired from the scanner and may be displayed in a monitor 303. The intensities may be processed to make genotype calls (i.e., determining the genotype based upon probe intensities) on the workstation 302. The intensities may be processed and stored in the workstation or in a data server 306. The workstation may be connected with the data server through a local area network (LAN), such as an Ethernet 305. A printer 304 may be connected directly to the workstation or to the Ethernet 305. The LAN may be connected to a wide area network (WAN), such as the Internet 308, via a gateway server 307 which may also serve as a

firewall between the WAN 308 and the LAN 305. In preferred embodiments, the workstation may communicate with outside data sources, such as the National Biotechnology Information Center, through the Internet. Various protocols, such as FTP and HTTP, may be used for data communication between the workstation and the outside data sources. Outside genetic data sources, such as the GenBank 310, are well known to those skilled in the art. An overview of GenBank and the National Center for Biotechnology information (NCBI) can be found in the web site of NCBI ~~(http://www.ncbi.nlm.nih.gov).~~

Please replace the paragraph on page 32 starting on line 3 with the following amended paragraph:

Primer design. After genes or genomic regions of interest were identified, PCR primers were designed in preparation for carrying out long PCR to produce amplicons ranging from 3 - 15 KB, using a variety of publicly and commercially available programs, i.e., Primer 3 ~~(www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi)~~ (available at the Whitehead Institute for Biomedical Research website), Amplify 1.2 (Engels et al. 1993), Oligo 6 (SR Lifescience,www.lifescience-software.com). Primers were tested on a pool of DNA produced from three different Coriell samples, cDNA or genomic DNA depending on the project.

Please replace the paragraph on page 40 starting on line 2 with the following amended paragraph:

Automation of sample preparation allowed a reduction in reagent volumes and reduced reagent costs by 33%. Automated array handling and analysis doubled the throughput possible. Ultimately, two skilled research assistants could routinely prepare sample, hybridize and analyze 40 arrays per day. Over the course of the two-year program all or part of 25,051 human genes (8.3 Mb) including some promoter regions were screened in 40 unrelated individuals of 3 different ethnic origins, producing a total of more than 15,000 SNPs which have been deposited in dbSNP (available at the NCBI website). ~~(http://www.ncbi.nlm.nih.gov/SNP).~~